

구분모드 합성법을 이용한 단백질 동역학의 전산해석 모델링

Computational Modeling with Component Mode Synthesis for Protein Dynamics

김재인 * · 엄길호 ** · 나성수 †
Jae In Kim, Kilho Eom, and Sungsoo Na

1. 서론

전통적으로 단백질의 동역학적 특성 이해에 있어서는 분자동역학이 사용되어왔다. 분자동역학은 이전의 연구에서 보여지듯이 단백질의 동역학적 특성을 이해하는데 좋은 결과를 제공하여 주었다. 그러나 분자동역학은 단백질의 모든 원자와 결합을 고려함으로써 컴퓨터로 계산하는 시간과 용량에 있어서 제한이 있게 된다. 단백질은 그 생물학적 기능을 하는데 있어서 마이크로초 단위의 시간을 필요로 하는데, 분자동역학은 나노초 단위의 결과를 보이게 된다. 이러한 것은 단백질의 동역학적 특성을 이해하는데 있어서 상당한 불편을 감수하게 하는 부분이다. 이러한 문제를 극복하기 위해 노말 모드 해석법과 탄성 네트워크 모델이 소개되었다. 탄성 네트워크 모델을 기반으로 가우시안 네트워크 모델이 제안되었다. 가우시안 네트워크 모델은 단백질 구조체를 단순한 알파 탄소의 1 차원 구조로 가정한다. 가우시안 네트워크 모델은 단백질 동역학의 노말 모드와 섭동의 빠른 계산을 수행하게 해주었다. 그럼에도 거대 단백질의 경우에는 몇가지의 비효율성이 나타났다. 이를 보완하기 위해 다른 방법들이 사용되었다. 부분 구조합성법이나 질량 압축법을 이용한 단백질의 해석이 이미 이루어졌고, 좋은 결과를 보여주었다. 이 논문에서는 구분모드 합성법을 이용하여 단백질의 동역학적 특성을 해석하였다. 반복적인 구조의 단백질이 아닌 일반적인 구조의 단백질에 적용하였고 단백질의 동역학적 특성을 잘 나타내었다.

2. 가우시안 네트워크 모델

가우시안 네트워크 모델은 단백질 동역학을 연구하기 위한 단순 질량-스프링 모델이다. 노드의 위치는 protein data bank(PDB)에서 얻은 실험 데이터를

바탕으로 결정하게 된다. 이런 단순화된 모델은 노드 간의 거리가 r_c 값 (7Å) 보다 작으면 하모닉 스프링으로 연결된다. 이러한 구조의 포텐셜 에너지는 다음과 같다.

$$P = \frac{\gamma}{2} \sum_{i,j}^N (r_{ij} - r_{ij}^0)^2 H(r_c - r_{ij}^0) \quad (1)$$

$H(r_c - r_{ij}^0)$ 는 헤비사이드 함수로 $r \geq 0$ 이면 $H(r_c - r_{ij}^0) = 1$, 다른 경우에는 $H(r_c - r_{ij}^0) = 0$ 이 된다.

또한 섭동 행렬 Q 는 다음과 같이 나타나게 된다.

$$Q_{ij} = \langle (r_i - \langle r_i \rangle) \cdot (r_j - \langle r_j \rangle) \rangle \quad (2)$$

3. 구분 모드 합성법

이 논문에서는 단백질의 해석을 위해 구조동역학에서 사용되는 구분 모드 합성법을 이용하였다. 구분모드 합성법은 거대 구조의 동역학적 문제를 풀기 위해 고안되었다. 구분 모드 합성법은 한 개의 주어진 구조체가 여러 개의 부분구조로 구성되어 있다고 가정한다. 각각의 부분구조에 대해 고유치 문제를 풀고 각각의 부분구조를 하나의 구조체로 구축하게 된다. 구분모드 합성법은 본래의 구조와 비교했을 때 더 작은 크기의 문제를 다룬다는 것에서 이점이 있다.

부분구조체 s 를 가정하고 임의의 점 P에서의 거리 벡터를 u_s 한다. 거리 벡터는 공간변수의 함수와 시간변수의 일반좌표계의 곱으로 나타나게 된다.

$$u_s(P, t) = \Psi_s(P) q_s(t) \quad (3)$$

부분구조의 키네틱 에너지는 일반좌표계를 이용하여 다음과 같이 나타낼 수 있다.

$$T_s = \frac{1}{2} \dot{q}_s^T M_s \dot{q}_s \quad (4)$$

M_s 은 부분구조의 질량 행렬이다.

이와 유사하게 포텐셜 에너지를 나타낼 수 있다.

$$V_s = \frac{1}{2} q_s^T K_s q_s \quad (5)$$

† 교신저자; 고려대학교 기계공학과

E-mail : nass@korea.ac.kr

Tel : (02) 3290-3854, Fax : (000) 000-0000

· 고려대학교 기계공학과

.. 한국과학기술연구원

K_s 는 부분구조의 강성 행렬이다.

여기서 M_s, K_s 는 고유벡터 Ψ_s 와 연관되어 대각 행렬로 나타낸다.

$$M_s = \Psi_s^T m_s \Psi_s, \quad K_s = \Psi_s^T k_s \Psi_s, \quad (6)$$

m_s 와 k_s 는 부분구조체 본래의 질량과 강성 행렬이다.

라그랑지의 식에 의해 운동방정식을 얻을 수 있다.

$$\frac{d}{dt} \left(\frac{\partial T_s}{\partial \dot{q}_s} \right) - \frac{\partial T_s}{\partial q_s} + \frac{\partial V}{\partial q_s} = Q \quad (7)$$

(4),(5)의 식을 (8)에 대입함으로써 다음의 식을 얻는다.

$$M_s \ddot{q}_s(t) + K_s q_s(t) = Q \quad (8)$$

여러 개의 부분구조체를 한 개의 구조체로 합하기 위해 $N \times N$ 블록 대각 행렬을 만들게 된다.

M_d, K_d 는 질량과 강성의 블록 대각 행렬을 나타낸다.

$$M_d \ddot{q}(t) + K_d q(t) = Q \quad (9)$$

부분구조의 모음에 대한 운동방정식은 다음과 같이 나타낸다.

일반좌표계의 벡터 $q(t)$ 는 부분구조간의 중복된 구속조건 값들을 포함하고 있다. N_c 개의 중복 구속 조건이 있다고 가정하고, $q(t)$ 가 M 차원을 가지고 있다고 하면, 독립변수의 개수는 $n = M - N_c$ 로 나타나게 된다. 여기에서 $q(t)$ 와 $\eta(t)$ 간의 좌표 변환을 다음과 같이 나타낼 수 있다.

$$q(t) = \delta \eta(t) \quad (10)$$

δ 가 구속 행렬이고 $\eta(t)$ 가 n -차원의 독립 일반 좌표계 벡터이다.

이렇게 새롭게 구성된 일반 좌표계에서 운동방정식은 다음과 같이 나타낸다.

$$M \ddot{\eta}(t) + K \eta(t) = Q \quad (11)$$

여기에서

$$M = \delta^T M_d \delta, \quad K = \delta^T K_d \delta \quad (12)$$

이 행렬들은 $n \times n$ 의 결함된 구조체의 행렬이다. 이 M 과 K 행렬을 이용해서 고유치 문제를 풀 수 있다.

$$K U = M U \lambda \quad (13)$$

4. 시뮬레이션 결과

해석을 위해서 헤모글로빈 (pdb code :1a3n)이 사용되었다. 헤모글로빈은 572개의 노드로 구성되고, 크게 4개의 부분으로 나눌 수 있다. 시뮬레이션에

서는 이를 바탕으로 4개의 부분구조로 나누었다.

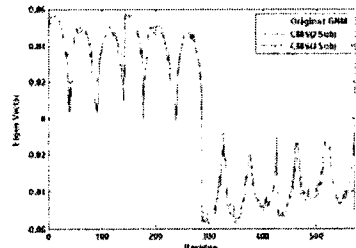


그림 1: 헤모글로빈에서의 가장 낮은 모드 비교

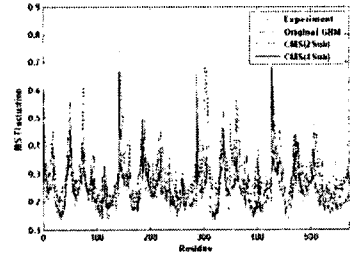


그림 2: 헤모글로빈의 섭동

그림 1에서는 고유벡터를 바탕으로 한 모드비교 결과를 확인 할 수 있다. 원래 구조를 해석한 것과 구분모드 합성법을 이용한 결과를 비교하였는데 헤모글로빈의 구조간 움직이는 방향을 알 수 있다. 그림 2에서는 열에 의한 섭동을 비교함으로써 실험치와 시뮬레이션 결과를 비교 할 수 있다. 그 값이나 경향성에 있어서 아주 유사한 결과를 보여주고 있다.

5. 결 론

구분모드합성법을 이용한 단백질 구조의 동역학적 특성 해석이 아주 잘 나타났다. 구분모드 합성법을 이용하여서 고 주파수 영역의 모드를 제거함으로써 기존의 방법으로 해석하기 힘들었던 보다 큰 크기의 단백질을 해석하는데 응용 가능할 것으로 보인다. 또한 이 방법을 기존의 탄성네트워크 모델로 해석이 불가능했던 초거대 분자 구조체에 도 사용하는 것도 가능할 것으로 보인다.

후 기

This work was supported by the Korea Science and Engineering Foundation (No.R11-2007-028-00000-0) and also acknowledges the support by Basic Research Program of the KOSEF under grant No. R01-2007-000-10497-0.